Contribution to the assessment of psychophysical properties of continuous text reading performance tests and contrast sensitivity tests

Brussee, T.

2018

SUMMARY AND GENERAL DISCUSSION

8

SUMMARY

Background and aims

This thesis contributes to understanding the psychophysical aspects of continuous text reading tests and contrast sensitivity (CS) tests. Both types of tests are important measurement tools and are used in ophthalmic and optometric clinical practice and research to investigate visual functioning. Reading performance and CS are strongly related to vision- and health-related quality of life and provide a broader understanding of visual performance.[1,2]

In our aging society, visual impairment poses an increasing burden on the lives of many older adults; both a decrease in health related quality of life and an increase in depression and/or anxiety disorders have been reported.[3,4] After accounting for population growth and ageing, the global prevalence of blindness decreased from 0.75% in 1990 to 0.48% in 2015. However, in absolute numbers blindness and visual impairment are increasing worldwide due to population growth, aging of the population, and the higher demands of visual performance.[5] In the Netherlands in the year 2008 an estimated 311,000 people were at least moderately visually impaired. This number is expected to have increased by 18% in 12 years which may lead to 367,000 visually impaired people in the year 2020.[6]

Reading is of great importance for participation in society and is one of the main reasons for visually impaired persons to seek low-vision rehabilitation.[7,8] Near vision assessment is increasingly included in epidemiological studies on blindness and visual impairment[5,9] and also in clinical studies and practice.[10,11] Especially in patients with age-related maculopathy (ARM) where central vision is affected, reading tests are more informative than letter charts using single optotypes to assess distance and near VA.[12-14] A reliable reading test allows an analysis of the effectiveness of low vision rehabilitation and optical needs, and provides a standardised measure of reading performance.[15,16]

In contrast to distance acuity, for reading performance tests there is no consensus regarding which test should be used. In 1988, the Visual Function Committee of the International Council of Ophthalmology (ICO) published a standard for reading charts to establish calibrated reading acuity measures.[17] However, a variety of (un-) standardised reading tests are still used in clinical practice and research.[18,19] This thesis focuses on continuous text reading tests for out loud reading, which are short duration tests that resemble daily reading material and are suitable for clinical practice and research. These reading tests allow a simultaneous measurement of reading parameters such as: reading acuity, critical print size, reading speed and reading mistakes. There are different other types of reading tests, e.g. mixed-contrast tests, comprehension tests, and tests with unrelated words. These tests all serve different purposes and are beyond the scope of this thesis.

There are several continuous text reading tests commercially available in the Dutch language: e.g. the Colenbrander Reading Card, the International Reading Speed Texts

(IReST), the Laboratory of Experimental Ophthalmology (LEO) test, 'de Nederlanders' (NED) (meaning 'the Dutchmen'), and the Radner Reading Charts. It seems evident that reading tests should be developed according to a standard,[20,21] however, of the tests that are available in Dutch, only the IReST and Radner Reading Charts have been psychophysically evaluated. It is unclear how standardised and unstandardised reading tests differ in relation to norm values, precision, agreement and feasibility of the tests.

CS describes the ability to see low-contrast patterns which can be impaired in many ophthalmic conditions.[22] CS is an important additional measure of visual function and independently associated with difficulties in everyday activities such as reading.[23-25] The USA National Research Council committee on disability determination, recommends to assess CS as a supplementary basis for individuals with visual impairments.[26] CS is a function, showing the relation between contrast thresholds and spatial frequencies. Instead of determining a whole CS function curve, contrast thresholds can also be determined at peak spatial frequency.[27] The Mars Letter Contrast Sensitivity Test is a small-format portable letter chart to measure spatial CS at peak sensitivity,[28] and has reasonable to good agreement with the Pelli-Robson chart.[29] Another, less widely used clinical test to measure CS, is the temporal CS test which measures temporal CS by presenting a flicker stimulus to the eye.[30] The concept of temporal CS is to better isolate neural function and has been shown to be useful in the detection and monitoring of several ocular diseases such as glaucoma, optic neuropathy, ARM and retinitis pigmentosa.[30-33] Recently a temporal CS test was developed[34] by making an adaptation to the C-Quant device.[35,36] Although it has been described that spatial CS plays an important role in reading performance, especially for persons with visual impairment,[27,37-40] the association between temporal CS and reading speed has not received attention yet. Comparison between the role of spatial CS (optical and neural) and temporal CS (solely neural) in the association with reading speed may provide insight into how reading performance is related to primary optical or neural or to both effects.

The goal of this thesis is to contribute to understanding the psychophysical properties of continuous text reading tests and different types of CS tests. In the first part, five continuous text reading tests available in the Dutch language were analysed. To this end, firstly, an overview of the literature on available continuous text reading tests was provided and the measurement properties of these tests were described. Secondly, the reliability and equality of sentences and paragraphs of five (un-) standardised reading tests available in the Dutch language were investigated in a sample of normally sighted adults of various ages. Thirdly, the norm scores, precision, agreement and feasibility of the five Dutch continuous text reading tests were evaluated in normally sighted and visually impaired participants. Fourthly, various visual and non-visual variables and their association with reading performance in normally sighted subjects were given. In the second part, the primary factors of CS and their association with reading performance was given. Firstly, a spatial and a temporal CS test and their associations with reading performance were analysed in normally sighted and visually impaired participants. Secondly, the precision of the spatial and temporal CS test and agreement between the two methods was assessed.

Part I: Continuous text reading performance tests

Systematic review of continuous text reading performance tests

Before selecting which reading test should be used in research or clinical practice, it is important to establish which reading tests have been developed to scientific standards and whether the measurement properties are properly evaluated. In this thesis, a literature search was performed to give an overview of available continuous text reading tests (**Chapter 2**). Systematic methods for selecting studies were used. In the available literature four types of data were extracted: 1) reading test characteristics, 2) design aspects of the studies, 3) algorithms for scoring reading performance and their reliability, and 4) general and reading test specific measurement properties. The measurement properties that were assessed were obtained from the Consensus-based standards for the selection of health measurement instruments (COSMIN).[41] The criteria were partly modified and some criteria were specifically added; content validity, internal consistency, reliability, cross-cultural validity, and generalisability were evaluated.

The search strategy identified 2334 articles, of whom 20 relevant articles were found on measurement properties of continuous text reading tests in various languages. Measurement properties for three of these tests were described in scientific studies; IReST, the Minnesota Low-vision Reading Test (MNread), and the Radner Reading Charts. Although other continuous text reading tests were found, no publications on measurement properties of these tests were available.

Considering reading test characteristics, the developers made different choices with regard to text type (paragraphs or single sentences), lexical difficulty, and word length. The IReST consists of 10 text paragraphs based on encyclopaedia material of about 141 words and with a text difficulty of grade 4 to 6 (age 10 to 12 years). The IReST was designed to measure reading performance and to compare various magnification modalities for longer text paragraphs and uses one print size. The MNread and the Radner Reading Charts adopted a logarithmic progression of print sizes and were designed for simultaneously measuring different reading parameters. The MNread consists of 12 word sentences displayed on three lines and with a 3rd grade (9 years) text difficulty. The Radner Reading Chart consists of 14 word sentences displayed on three lines and with a 3rd to 4th grade (9 to 10 years) text difficulty. Considering design aspects of scientific studies, only half of the studies provided comprehensive information on patient characteristics. All studies (except for Patel et al.[42]) have been performed under standardised (laboratory) conditions.

Considering algorithms for scoring reading performance, the reading parameters that were described were: average or mean reading speed, maximum reading speed (MRS), reading acuity (RA), and critical print size (CPS). For every parameter, different scoring rules and terminology were used in the studies. For example for the average reading speed two different scoring rules and for the CPS six different scoring rules were described. Furthermore, of the thirteen studies in which the precision of the reading test was assessed, only eight used the limits of agreement methodology as originally proposed by Bland and Altman,[43] even though this is the method of choice for evaluating agreement (comparison of clinical tests) and precision (repeatability or reproducibility) for clinical tests.[44]

Considering general and reading test specific measurement properties, all three reading tests scored high on content validity, however content validity was not described for every language in which the MNread and Radner Reading Charts were commercially available. The IReST scored best on inter-language comparison. The MNread scored well in repeatability studies (repeated measurements under the same conditions). The Radner Reading Charts had good inter-chart reliability and reproducibility results, and a well-developed rater system. The MNread and the Radner Reading Charts were tested both in normally sighted populations and in subjects with visual impairment. Studies investigating the IReST were performed only in a normally sighted population.

Differences in the design of the MNread and the Radner Reading Charts should be considered. Even though the tests are very similar in the way they provide measures of reading speed, RA and CPS, there is a difference between these tests concerning standardisation of sentences. The MNread is less standardised in terms of syntactical and geometric structure (number of syllables, characters, position and word length) compared to the Radner Reading Charts. However, there are no 'direct comparison' studies between the MNread and the Radner Reading Charts available, making it unclear whether these differences in study design will lead to better reliability results of the Radner Reading Charts. This issue has also been discussed by the authors.[45,46] The design and purpose of the IReST is very different in comparison with the MNread and Radner Reading Charts. The paragraphs are available in one print size and can often only be used with magnifier aids in subjects with visual impairment. Longer text paragraphs may provide a supplemental tool in reading diagnostics, such as for highlighting difficulties with prolonged reading, standardised print size and navigation issues. It has been reported that reading speed with single sentences is similar to longer paragraphs, but that single sentences may overestimate reading speed in prolonged reading situations.[12,15,47,48] In turn, it has also been suggested that reading speed measures are more reliable for longer paragraphs than for single sentences (e.g. MNread, Radner Reading Charts) due to lower variance.[49,50]

The quality and availability of research to assess measurement properties of reading tests varied and some important measurement properties were missing.

The IReST has good trans-linguistic comparability and scores well on cross-cultural validation in 17 different languages, and is explicit about the translation process. Direct comparison between languages for the Radner Reading Chart and the MNread needs further research. On the other hand, in clinical studies the main focus is often on relative improvement of the reading acuity or reading speed before and after an intervention. This relative improvement can be compared if a reading test has proven to be reliable within one language. Studies on the reproducibility of the IReST and the MNread tests were missing, whereas the Radner Reading Chart had good reproducibility studies. On content validity all three reading tests scored high. However, for both the MNread and the Radner Reading Chart, content validity studies were not available for all languages. For the development of reading tests in new languages, we recommend following the definition rules of the authors of the abovementioned tests on content validation, to have the translation performed by linguistic experts, and at least, to perform analyses for internal consistency and for the selection of sentences or paragraphs.

In conclusion, an important consideration regarding which reading test to choose is whether the tests have been developed and tested according to scientific standards. For standardised reading tests, print size should be the only parameter affecting performance throughout the chart. Therefore, sentences with a high comparability are preferred, which could possibly be acquired best through extensive standardisation of the sentences. Finally, standardised (laboratory) test conditions are preferred for studies on development and reliability testing of reading tests. This will make comparison between studies possible.

Reliability and equality of sentences and paragraphs

In standardised logarithmic reading tests, print size should be the only parameter affecting reading performance throughout the chart. This means that all sentences should be equally difficult to read and that all sentences should require the same reading time and have similar (occasional) mistakes when presented in the same print size. An important step in the design of a reading test is to investigate sentence or paragraph reliability and to test whether these are equally difficult. This step is often not described in the literature.

Therefore in Chapter 3 the content validity (reliability and difficulty of the sentences and paragraphs) in addition to the internal consistency (test reliability) within and between reading tests was studied in a group of normally sighted subjects. The aims of this chapter were: 1) to provide norm values of reading speed and mistakes, 2) to investigate sentence and paragraph difficulty and reliability of each test, 3) to assess agreement between the reading tests. The reading tests available in the Dutch language that were studied were: The Colenbrander Reading Card,[21] the IReST,[49,50] the LEO,[51] NED,[18] and the Radner Reading Charts.[48,52] The reading test characteristic are described in more detail in Table 1.

Table 1. Reading tests characteristics

Reading Test	Print size	Range in logMAR at 40 cm	No. of charts
Colenbrander Reading Card (Colenbrander 1990s, USA)	Logarithmic progression print size	1.2 to -0.1 logMAR	1 chart
IReST (Trauzettel 2005, Germany)	One print size	0.4 logMAR	10 charts
LEO (Kooijman and Beerthuizen 1996, Netherlands)	Logarithmic progression print size	1.4 to -0.3 logMAR	9 charts
'de Nederlanders' (Snellen 1862, Netherlands)	Logarithmic progression print size	0.5 to 0.1 logMAR	1 chart
Radner Reading Charts (Radner 1998, Austria)	Logarithmic progression print size	1.2 to -0.2 logMAR	3 charts

a: mean number of words are given as there is a difference between sentence or paragraph length, No.: number

In this study, 71 normally sighted persons (30 males, 41 females) mean age 55 years (SD 20; range 18 to 86) with a binocular distance acuity of logMAR 0.20 (Snellen 0.63) or better were evaluated. The mean education level was 12 years (SD 3; range 6 to 16), which is comparable to the 12th grade educational level. For comparability purposes, all sentences and paragraphs available in the different reading tests studied were printed (one on each page) in Times New Roman type size 12 point. The order of the five tests was randomised; also, for each test, 10 versions were made with randomised sentence or paragraph order, to spread any possible practice or fatigue effects. All testing procedures were monitored with video and audio recording to measure reading speed and reading mistakes afterwards. To compare the five tests, each with different average word lengths, adjustments had to be made. Consequently, the number of characters and standard words per minute were calculated.[53] To test whether sentence and paragraph difficulty was equal between the sentences or paragraphs used within each reading test, mean reading speed in standard words per minute for each sentence and paragraph were compared to the mean reading speed within a test. In addition, the 95% limits of agreement for different reading tests (for both reading speed and mistakes) was analysed.

Reading speed was highest for the Radner Reading Chart (179 standard words per minute) and lowest for NED (142 standard words per minute). The least reading mistakes were made with the Radner Reading Chart and the most reading mistakes were made with the NED and Colenbrander Reading Card. Of all the sentences and paragraphs read, no reading mistakes were made in 75% of the cases. Although the interrelatedness (internal consistency) for all tests was high (Cronbach's alpha

No. of sentences or paragraph(s) per chart	No. of words per sentence or paragraph	Average word length (characters)	Lexical difficulty
14 sentences	20[a]	3.42	4th grade (10 years)
1 paragraph	141[a]	4.85	4th to 6th grade (10-12 years)
18 sentences (55 total)	30[a]	4.84	4thgrade (10 years)
5 paragraphs	61[a]	4.99	Literary level
14 sentences (24 total)	14	5.07	3rd to 4th grade (9-10 years)

coefficient of ≥0.983), the sentences and paragraphs were not equally difficult. On the Colenbrander Reading Card 7/24 sentences were read significantly faster vs. 5/24 read slower, sentence reliability [0.56-0.87]; IReST 3/10 vs. 3/10 [0.94-0.97]; LEO 14/55 vs.15/55 [0.64-0.92]; NED 2/6 vs. 3/6 [0.83-0.94]; Radner Reading Charts 4/24 vs. 3/24 [0.73-0.87]. The highest agreement was found between the Colenbrander Reading Card and IReST (1.2 standard words per minute) and the Colenbrander Reading Card and the NED (0.01 mistakes per 100 characters). The least agreement was found between the NED and the Radner Reading Chart (36.4 standard words per minute) and the Colenbrander Reading Card and the Radner Reading Chart (0.15 mistakes per 100 characters).

For all tests, several sentences or paragraphs differed significantly from the mean; sentence and paragraph reliability varied both within and between reading tests. The influence of text properties such as crowding, font type and print size on this inequality can be excluded, as these were equalised for all reading tests. However text content (e.g. linguistic difficulty, syntactical structure and contextual cues) differed between the tests, and influenced the results. For example the higher level of linguistic difficulty of the NED, may explain lower reading speeds, more reading mistakes and the lowest equality of paragraphs. Also less agreement between the NED and other tests was observed concerning reading speed and mistakes. The high level of standardisation of the Radner Reading Charts in terms of lexical difficulty, syntactical structure and geometric proportion, might explain the highest number of equally difficult sentences and the low agreement when compared to the other tests.

In summary, the sentences of the Radner Reading Chart are appropriate to measure reading acuity as well as reading speed in a heterogeneous population. Furthermore, it was shown that the design principles in terms of lexical difficulty, syntactical structure and geometric proportion of the Radner Reading Charts led to high equality of the sentences on the chart in terms of difficulty, as well as to the highest reading speed and lowest number of reading mistakes. The design principles[48,52] of the test are a good example for the design of new reading tests. The higher number of sentences or paragraphs that differ significantly from the mean for the Colenbrander Reading Card, LEO and NED might lead to the consideration of changing or choosing specific sentences or paragraphs on the existing reading charts when used for research purposes. To measure reading speed using longer text paragraphs, the IReST had higher paragraph equality and reliability compared with the NED. However, since the reliability of the NED is low, the validity of comparing the IReST with the NED can be questioned. Therefore further research on the reliability of the IReST is needed.

Measurement properties of continuous text reading tests

In the former chapter, reliability and equality of the sentences and paragraphs of the reading tests were investigated in normally sighted adults using the same font and print size. In **Chapter 4**, the psychophysical properties of the Colenbrander Reading Card,[21] the IReST,[49,50] the LEO,[51] the NED,[18] and the Radner Reading Charts[48,52] were evaluated using their original designs in patients with maculopathies and in students with normal sight. The aims of this chapter were to 1) provide norm scores, 2) compare the inter-chart and inter-session reliability of the reading tests, 3) investigate agreement between tests, and 4) evaluate feasibility of the tests for different groups of participants.

In this study, we evaluated 25 normally sighted students (4 males, 21 females) with a mean binocular distance acuity of logMAR -0.1 (Snellen 1.25) and 47 patients with maculopathies (21 males, 26 females) with a mean monocular distance acuity of logMAR 0.58 (Snellen 0.25). Of the 47 patients, 43 had ARM, 2 had Best disease, 1 had Stargardt maculopathy, and 1 status after macular hole successfully treated with pars plana vitrectomy. Patients were asked to read (monocularly) each sentence and paragraph aloud as quickly and as accurately as possible at a preferred viewing distance ranging from 22 to 40 cm. The order of the logarithmic tests and (if applicable) of the charts was randomized, the IReST texts were read last. To be able to test the reliability of the IReST, enlargement of the texts was needed. The procedures for the students were almost the same, however tests were read binocularly at a fixed reading distance (40 cm). To determine the precision of the reading tests, the whole testing procedure was repeated within a period of 2-10 weeks. Assessment of the reproducibility of the chart (inter-chart reliability) and test (test-retest reliability) of the reading performance parameters were performed. To compare methods between the five reading tests, agreement between reading parameters (MRS, RA, CPS) between the reading tests was analysed by determining the 95% limits of agreement. Feasibility of the reading tests was evaluated by describing whether the layout of the reading tests allowed an actual measurement in participants with or without visual impairment.

On the different charts, norm values for the MRS in words/min were 82 to 141 for patients and 164 to 237 for students. For RA in logMAR the norm values were 0.44 to 0.69 for patients and -0.09 to -0.05 for students. For CPS in logMAR the norm values were 0.47 to 0.84 for patients and 0.08 to 0.11 for students. Regarding the precision of the reading tests, in most cases the differences in CR between the tests for RA and CPS were within the limit of one line (0.1 logMAR). However, important to note is that, in the patient group, the inter-session CR for CPS for the Radner Reading Charts was 0.10 and 0.11 logMAR lower compared with the CR of the LEO and Colenbrander Reading Card, respectively. Therefore, measuring the CPS for patients with maculopathies might be most reliable with the Radner Reading Charts. The present study shows that the agreement between different reading tests can differ for different participant groups. For example, agreement between the Colenbrander Reading Card, LEO and Radner Reading Charts was high for the students, whereas for the patients some significant differences in agreement between the reading tests were found. For the patients, the highest agreement was found between the LEO and Radner Reading Charts for both RA and CPS; the IReST showed the highest agreement with the LEO for measuring MRS.

The value of the NED for use in clinical practice can be argued, as only 57% of the patient group was able to read the chart. This can be explained by the availability of print sizes on the chart. The largest print size of the NED corresponds with an RA of 0.50 logMAR (Snellen 0.32) at 40 cm, whereas the mean distance VA of the patients was 0.65 logMAR (Snellen 0.22). This in combination with the linguistic difficulty of the language may explain why almost half of the patient group was not able to read the NED. In addition, for the students it was not possible to determine RA and CPS with the NED due to ceiling effects. Also the use of the Colenbrander Reading Card for a young and normally sighted population is debatable. For the Colenbrander Reading Card ceiling effects were found for RA, because 21 of the 25 students could read the smallest sentence on the Card. This ceiling effect can probably also contribute to the lower inter-session CR for RA for the Colenbrander Reading Card found compared with the LEO and Radner Reading Charts. Remarkable is also that, in the patients, the RA and CPS for the Colenbrander Reading Card were up to 0.13 logMAR and 0.11 logMAR better, respectively, and higher reading speeds were found. An explanation might be the short words used on the Colenbrander Card (i.e. easier to read), leading to higher reading speeds and more contextual cues which could lead to guessing and result in overestimations of RA and CPS.

In conclusion, it was shown that reading performance results obtained with function-based reading tests are not always reliable and that reading parameters could not always be properly assessed in patients with maculopathies and in students with normal sight. Therefore, choices regarding which reading test to use in clinical practice, and especially for research purposes, should be based on both the feasibility and reliability of the reading test. Both the Radner Reading Charts and LEO showed good reliability results and the best agreement was found between these two tests. The Radner Reading Charts are available in several European languages, whereas the LEO is only available in the Dutch language. This makes the Radner Reading Charts more suitable for research studies across different languages in order to allow standardised and comparable analyses of reading performance. In addition, it can be concluded that measuring the CPS for patients with maculopathies might be most reliable with the Radner Reading Charts as well. The longer text paragraphs of the IReST - which can be a supplemental tool for providing information on reading performance over prolonged reading - needs further investigation as this was the first reliability study of the IReST in a group of visually impaired participants.

Visual and non-visual factors and the association with reading performance

Several visual parameters have been reported to influence the reading process, i.e. VA, CS, field of view (scotomata), where VA only slightly accounts for the variability in reading speed.[38,54-57] The aims of this study described in **Chapter 5** were 1) to assess how non-visual variables (age, education level, reading habits) and visual variables (VA, CS, and straylight) affect reading performance in terms of reading speed and mistakes, and 2) whether prolonged reading influences reading performance with single sentences and paragraphs.

In this study, 71 normally sighted subjects (30 males, 41 females) mean age 55 years (SD 20; range 18 to 86) participated, with mean binocular distance VA of logMAR -0.01 (Snellen 1.02). The mean education level was 12 years (SD 3; range 6 to 16), which is comparable to the 12th grade educational level. All available IReST paragraphs and Radner Reading Chart sentences were printed (one on each page) 12 point Times New Roman. Tests were performed binocularly and participants were asked to read one page after the other, reading each sentence and paragraph aloud. All testing procedures were monitored by video and audio recording. These recordings were used to later measure reading speed and reading mistakes. CS was measured with the Mars Letter CS Test[28] at a test distance of 40 cm.

For both the IReST paragraphs and the Radner Reading Chart sentences, the correlation between independent variables and reading speed was moderate to strong for age, educational level, near and distance VA, and CS. Multiple regression analyses showed that CS was independently associated with reading speed measured with the IReST paragraphs (P=0.002). The finding was confirmed for the Radner Reading Chart sentences (P=0.021). There was also an interaction effect between age and education in the association with reading speed. For the older age groups, reading speed was less influenced by education level compared with the younger age groups.

Distance and near VA seemed independently associated with reading speed but when other variables were taken into account, these associations were no longer significant. To gain insight into prolonged reading and the potential fatigue process, at each measurement, reading speed was plotted against test time. The linear mixed effect model showed that, after correction for relevant variables, the association between point in time and reading speed was not significant over a period of about 23 minutes, meaning that reading speed remained practically stable over this period of time.

The importance of measuring CS in clinical practice and research should be discussed. The significant correlation found between distance VA and CS (r= -0.414; $P<0.01$) is particularly of interest, as the normally sighted group of participants investigated was a group with limited variability in CS and VA. It is also interesting that, although the correlation between CS and age was high (r= 0.60; $P<0.01$), CS was independently associated with reading speed, whereas distance and near VA were not significant in the association with reading speed. When assessing reading performance (for normally sighted subjects), especially if reading speed is lower than would be expected based on VA, measuring CS is of clinical importance.

Interesting to note is that the interaction effect between age and education showed that for the older age groups, reading speed was less influenced by education level compared with the younger age groups. An explanation for this might be the literacy level which can be enhanced throughout lifespan. In fact, the years of education may not accurately reflect a person's actual literacy level, especially in the older age groups. The present study also found an age-related decline in reading speed. This decline in reading performance may be due to subtle age-related visual deficits in healthy aging eyes, which might be the result of a combination of optical changes and changes in neural transmission.[58] Other, non-visual factors explaining the decline in reading performance in the elderly include the decline in cognitive mechanisms, such as working memory and information processing speed.[59,60]

The standardised tests used in the present study can be confidently used in clinical practice to reliably assess reading performance, as it was shown that when using these tests in normally sighted subjects, reading fatigue proved not to be an issue for reading durations up to about 23 minutes. However, this might be different for subjects with visual impairment. For different ocular health states, it seems that different durations of tests are needed to reveal a decline in reading speed due to reading fatigue.[61-63]

In conclusion, it was shown that CS was significantly associated with reading speed and an interaction effect was found between age and education. It is recommended to administer standardised reading speed and CS tests more often in routine clinical practice, if only to have a baseline value allowing to follow patients over time, especially in the elderly population where CS may decline before a loss of VA occurs. It is also recommended to assess CS in research and take age and education into account, especially for the interpretation and comparison of results between individuals and different studies.

Part II: Contrast sensitivity tests

Background and methods

For patients with macular pathologies, besides a decrease of VA also a loss of spatial and temporal CS has been reported.[30,64] Although spatial CS is associated with VA, it provides additional information on visual performance.[24] The strong association between CS and functional abilities such as face recognition,[65] driving,[66,67] and reading,[25] provide a rationale for including CS measurements in clinical practice and research related to functional vision and quality of life.[22]

In general, two aspects of CS can be measured: spatial CS and temporal CS. Spatial CS can be measured with gratings with a sample of spatial frequencies or with large letters that test peak spatial frequency. Spatial CS is influenced by the optics of the eye, even at lower spatial frequencies.[68] To measure flicker sensitivity or temporal CS, a flicker stimulus is presented to the eye. The concept of temporal CS is to better isolate neural function, as peak temporal CS is insensitive to optical blur.[69]

Spatial CS loss plays a critical role in reading, especially in persons with visual impairment,[70] but less is known about the association between temporal CS and reading speed. Both VA and spatial CS can be considered to originate from two more basic aspects of the human eye: the optical and the neural aspect. Temporal CS, however, targets the neural aspect only, and can be considered to underlie VA and spatial CS. It is possible though that the retinal mechanisms targeted by temporal CS do not precisely coincide with those that are important for VA and spatial CS. More insight into primary visual factors influencing reading performance is important for understanding a reduction in reading capacity in visually impaired patients with macular disease.

In this thesis the role of spatial CS (optical and neural) and temporal CS (solely neural) in the association with reading speed was evaluated to provide more insight into how visual tasks (such as reading) are related to primary optical or neural or both effects in a clinical sample of normally sighted adults of various ages (**Chapter 6**) and in visually impaired patients with macular disease (**Chapter 7**). In both of these chapters, 1) the associations between either spatial or temporal CS and reading speed was analysed, 2) the precision of the spatial and temporal CS test was assessed, and 3) agreement between the tests was determined.

For assessing spatial CS the Mars Letter CS Test was used. Regarding temporal CS tests, the C-Quant device with an adaptation to measure temporal CS was used.[34] The Radner Reading Charts were used to measure reading speed. The Mars Letter CS Test is a hand-held chart with three versions, each with a different sequence of 48 out of ten Sloan letters arranged in eight rows of six letters each. The contrast range corresponds to 0.04 to 1.92 log CS, with each successive letter representing an increment of 0.04 log units. Participants were asked to read one chart for each eye. Charts were presented in a randomized order and were presented on a reading stand. Chart luminance was around 90 cd/m^2. The temporal CS test implemented in the C-Quant device, consists of a central test field with a diameter of 3.3 degrees divided in two halves surrounded by areas of constant luminance. During the examination, a flicker stimulus was randomly presented to one of the halves of the test area with a maximum of 6 seconds and the participant had to press the left or right button to indicate in which half of the test field the flicker stimulus was present.

The first study (Chapter 6) evaluated 71 normally sighted persons with good binocular VA. Mean logMAR VA of the participants best eye was -0.01 (Snellen 1.02). To investigate the association between spatial CS and temporal CS and reading speed, Pearson's correlation analyses were performed. The correlations between spatial CS or temporal CS and reading speed and age were also investigated by presenting partial correlations, as these variables are highly correlated. The precision of the measurement was defined as the degree to which the same method produces the same result for repeated measurements in the same individual on the same day (repeatability). The session was repeated after 30 to 40 minutes.

The second study (Chapter 7) evaluated 47 participants with maculopathies, mean monocular distance VA was logMAR 0.58 (Snellen 0.25). To investigate the association between spatial and temporal CS and reading speed, first correlations between visual characteristics (distance VA, central loss, metamorphopsia, media condition) and non-visual characteristics (age and education) were analysed with Pearson's and Spearman's correlations. Linear regression models were fit to investigate the associations between either spatial or temporal CS and reading speed in uncorrected and corrected models. Adjustments were made for relevant visual confounders and patient characteristics. To determine the precision of the measurement the coefficients of repeatability and reproducibility were calculated for the spatial CS test and the temporal CS test. Measurements were repeated under the same conditions after 30 to 40 minutes and after a period of 2 to 10 weeks. In both studies the 95% limits of agreement between spatial CS and temporal CS were determined using Bland-Altman analysis.

8

The role of primary factors in CS and associations with reading performance

In the clinical sample of normally sighted adults a significant correlation was found between reading speed and both spatial CS (r= 0.470; $P<0.001$) and temporal CS (r= 0.258; $P=0.042$); partial correlations controlled for age were r= 0.175 ($P=0.166$) and r= 0.152 ($P=0.239$), respectively. Spatial CS and temporal CS decrease with increasing age. The correlations found between both spatial CS or temporal CS and reading speed, indicate that not only optical aspects but also neural aspects may be important in defining reading speed. However, partial correlations between reading speed and spatial CS or temporal CS controlled for age were no longer significant and especially for spatial CS found to be substantially lower than the original correlation. By correcting for age in a normally sighted group of participants with limited variability in CS, variability was reduced even further, explaining why the partial correlation between reading speed and spatial CS or temporal CS was no longer significant.

In the clinical sample with patients with maculopathies, reading speed correlated with both spatial CS (r= 0.35, *P*=0.015) and temporal CS (r= 0.66, *P*<0.001). Reading speed increased with increasing spatial and temporal CS. The temporal CS could not be performed by 10 participants because they were unable to see the test field and/or had delayed responsiveness. Participants unable to perform the temporal CS test were significantly older. In the uncorrected model spatial CS explained 12% of the variance and temporal CS explained 44% of the variance. The high correlation and higher explained variance between temporal CS and reading speed might be explained by the finding that temporal CS[71] as well as reading speed[72] are relatively unaffected (or less affected) by media opacity, but are affected by retinal disease. Based on relevant and significant correlations, the regression models were adjusted for distance VA, central loss and education level. The corrected models explained 64% and 61% of the variance for respectively spatial CS and temporal CS in the association with reading speed. However, the association between temporal CS and reading speed was not significant anymore. This can probably be explained by the high number of patients who had advanced geographic atrophy, which had a significant impact on distance VA and reading speed.

Precision of a spatial and temporal CS test

The precision of both the Mars Letter CS Test for measuring spatial CS and temporal CS with the C-Quant device were assessed. In the sample of normally sighted adults the precision of the spatial CS and temporal CS tests assessed with the coefficient of repeatability were 0.24 log units for the temporal CS test and 0.14 log units for the spatial CS test. Considering the precision of the tests in the visually impaired patients at the first administration measurement and at the retest, the coefficient of repeatability and reproducibility was 0.22 and 0.28 log units for the spatial CS test and was 0.33 and 0.35 log units for the temporal CS test, respectively. The difference in coefficients of repeatability between the two methods might be attributed to the psychometric differences between the methods. The spatial CS test, which makes use of 10 Sloan letters, has a 10 alternative forced-choice procedure compared with a 2 alternative forced-choice procedure of the temporal CS test. Furthermore, the coefficient of repeatability was higher in the patient group with maculopathies, which is in agreement with other studies where the coefficient of repeatability for visual parameters has been found to be higher for visually impaired subjects compared to normally sighted subjects.

Agreement between spatial and temporal CS

There is an essential difference between the two methods: spatial CS relates to a localised difference in contrast (letter in a surrounding) and temporal CS relates to a periodical difference in contrast (intensity changes back and forth over time). Moreover because the spatial CS test was a non-periodic test, it is defined by the Weber contrast formula and temporal CS is defined by the Michelson contrast formula. It is important to note that due to the different formulas used to determine contrast, there is a factor of approximately two difference between spatial CS and temporal CS resulting in a difference of 0.3 log units higher results for the temporal CS test. Considering agreement between the two methods, after applying the correction factor of 0.3 log units, the values for the temporal CS were 0.10 and 0.13 log units (normally sighted participants and participants with maculopathies, respectively) higher than those for spatial CS. The difference can potentially be attributed to optical mechanisms leading to a worsened image on the retina for the spatial (specifically, letter) CS.

Conclusion

The aim of **Chapter 6** was to gain insight into the association between spatial CS and temporal CS and reading speed in a sample of normally sighted adults of various ages. The significant correlation found between temporal CS and reading speed indicates that, besides optical components, neural aspects may also be important in defining reading speed. The stronger correlation between spatial CS and reading speed is suggested to reflect a deterioration of both optical and neural factors with increasing age. Concluding for this normally sighted group of participants, the influence on reading speed seems to be mainly optical in origin and, therefore, related to age effects. However, it cannot be excluded that also other factors (e.g. cognitive and visual) change with age and can partially explain the age-related decline in reading speed.

The aim of **Chapter 7** was to gain insight into the extent in which optical versus neural aspects influence reading speed for individuals with macular pathologies. For spatial CS and temporal CS moderate to strong correlations were found, respectively, with reading speed. The stronger association between temporal CS and reading speed is suggested to reflect a high sensitivity for neural integrity of the temporal CS test. However, it cannot be ruled out that other visual functioning tests or patient characteristics mask the association in a sample of patients with progressive ARM. The correlation between temporal CS and reading was higher than in the study in participants without macular pathologies. This might be explained by the larger variability in macular function in the present study, which seems to be strongly indicated by the temporal CS test.

Another aim of the two studies was to assess the precision and agreement of two relatively new tests for measuring spatial and temporal CS. Both the Mars letter CS test and the temporal CS with the C-Quant device were shown to be a reliable test for measuring spatial CS and temporal CS, respectively. A downside of the temporal CS test is that not all participants with severe maculopathies could manage to perform the test due to profound vision loss. On the other hand, the test might be promising as a diagnostic test to detect and monitor the progression of early ARM and to assess the efficacy of interventions.[73,74]

GENERAL DISCUSSION

Part I: Continuous text reading performance tests

Recommendations for future research

Based on the outcomes of this thesis, different suggestions can be made for future development and reliability assessment studies of reading tests. Improving the quality and content of study designs and methodology used, makes it easier to compare the outcomes between reading test studies. Regarding study design, including careful description of patient characteristics, use of objective and subjective lighting levels, good control of working distance, documentation of the number of raters and their training, careful documentation of scoring rules and the use of Bland-Altman analyses or similar for reproducibility and repeatability studies are recommended. For the development of reading tests in new languages, it is recommend to have the translation performed by linguistic experts, to follow the definition rules of the author of the test on content validation, perform a measurement for the selection of sentences, and performing an internal consistency analyses. The development of tests in different languages is much more complicated than simply translating phrases.

Regarding scoring algorithms for reading parameters, there is no consensus on the method of choice. There is less discussion on RA definitions, but MRS and CPS are not that clear. Consensus on this issue may facilitate the adoption of reading performance as a proper outcome measure in clinical settings and research. For clinical practice more subjective algorithms are suggested to be feasible, whereas in scientific research more objective algorithms are preferred. The repeatability of reading parameters is worse for visually impaired individuals, as the parameters may be difficult to determine and/or the reading speed versus print size curve might be fluctuating. In addition MRS does not allow differentiation of individuals who can read small print sizes at a given speed from those who can only read larger sizes at the same speed.[75] Using group data masks the variability within individuals; in some cases it is important to determine the patient's own criterion for change.[76,77] For this purpose, Calabrese et al.[78] recently introduced a fourth parameter: the Reading Accessibility Index (ACC). The index captures an individual's range of accessible print sizes and reading fluency within this range in a single value, without the need for curve fitting. Changes in an individual's ACC might be used to evaluate the effect of ophthalmic treatment, reading devices or reading rehabilitation programs. It could also be useful in research related to quality of life instruments and visual functioning indices. Further research could focus on which scoring algorithms for the different parameters are most reliable in different patient groups and therefore suitable for scientific research.

This thesis was the first to describe the content validity, internal consistency and reliability of the Colenbrander Reading Card, the LEO and the NED in a group of participants with normal sight and with maculopathies. By providing important patient characteristics such as age, reading habits, educational level, visual acuity, and if applicable, eye disease and severity, this gives an impression to which study population the results can be generalised. Studies in which the content validity and reliability were measured (mainly for a young and homogenous group of students) were available for the IReST and the Radner Reading Charts. For future research it would be interesting to evaluate cross-cultural validity for the different reading tests, as this was not part of this thesis and not described in the literature before, except for the IReST. For example, for direct comparison of absolute reading speed between languages a correction factor can be calculated. On the other hand, in clinical studies the focus is on relative improvement of reading parameters before and after an intervention. This relative improvement can be compared between studies in different languages, if a reading test has proven to be reliable within a language. Further reliability and feasibility studies especially among visually impaired participants with different eye pathologies would be interesting for the IReST. In this thesis, the reproducibility of the IReST was described for the first time for a group of normally sighted subjects and for a group of subjects with maculopathies.

Implications for clinical practice and research

One of the aims of the work presented in this thesis was to gain insight in the psychophysical properties of reading tests and to evaluate which reading test(s) should be used in clinical practice and scientific research. First the content validity (reliability and difficulty of the sentences and paragraphs) of the reading test was evaluated for normally sighted participants. Secondly, the psychophysical properties of the reading tests by their original design were evaluated for normally sighted subjects and with maculopathies. It was found that the Colenbrander Reading Card had a low percentage of sentence equality, and adjustments of the chart or choosing specific sentences needs to be considered if they were to be used for research purposes. For the Colenbrander Reading Card ceiling effects were found for RA, for normally sighted participants. The Colenbrander Reading Card was found to be less reliable in determining reading parameters compared to the LEO and Radner Reading Charts. For these reasons the use of the Colenbrander Reading Card is debatable.

Regarding the NED, it was found that the linguistic difficulty of the test resulted in lower reading speeds and a higher number of reading mistakes. For the normally sighted participants it was not possible to determine RA and CPS with the NED due to ceiling effects. Furthermore, only 57% of participants with maculopathies were able to read the NED. Moreover, unguided development of some versions of the NED has led to differences compared with its original design.[18,79] For these reasons it is recommended that the NED is no longer used for clinical practice or scientific research.

The LEO had a low percentage of sentence equality, and adjustments of the charts or choosing specific sentences need to be considered if they were to be used for research purposes. On the other hand the LEO scored good inter-session reliability results. Furthermore the agreement between the LEO and Radner Reading Charts was high. For the last two reasons the LEO (even though not the first method of choice) can be considered an option for the Dutch clinical practice.

The Radner Reading Charts had the highest percentage of equal sentence difficulties, the highest reading speed and lowest number of reading mistakes. Overall the Radner Reading Charts scored good inter-session reliability results. Furthermore, measuring the CPS for patients with maculopathies was most reliable with the Radner Reading Charts. Considering all psychophysical properties studied, the Radner Reading Charts are most appropriate for clinical and research purposes for measuring reading parameters.

The longer text paragraphs of the IReST can be a supplemental tool for providing information on reading performance over prolonged reading. The IReST had higher paragraph equality and reliability and almost equal reliability results compared to the NED. However, since the reliability of the NED is low, the validity of comparing the IReST with the NED can be questioned. Therefore for further research on the reliability of the IReST it would be interesting to compare this test with more elaborate standardised longer paragraphs such as recently developed by Radner et al.[80] The authors found that for a normally sighted group of subjects significant differences can appear between long paragraphs, despite good reliability, validity and high inter-item correlation of the paragraphs.[80] Considering the many parameters (text properties and content and psychophysical aspects of the reader) that can influence reading performance, there is probably no single definitive reading speed for an individual. These effects can be more profound for longer paragraphs on inter-session testing and it would be interesting to investigate the influence of psychophysical aspects of participants on reading performance in future research.

The future will most probably lay in the development of electronic reading devices for measuring reading performance.[81,82] However, for every newly developed reading test content validity and internal consistency of the sentences presented on the electronic devices will remain to be important, to lead to high equality of the sentences as well as to optimal reading speed and low number of reading mistakes. In the development of new electronic devices for measuring reading performance the use of sentences of already existing printed reading tests that have proven to be reliable can be considered. The advantage of electronic reading devices is that they can control several conditions, for example the reading distance, which is one of the greatest challenges and one of the most important factors in the testing method. A recently developed device is the Salzburg Reading Desk, in which reading distance and reading speed is measured by the device.[83] Another device is a novel instrument for logging reading distances which is measured with a ultrasonic transceiver technique.[84] Recently the English version of the MNread test was implemented on an iPad3, which yielded similar results to the chart version of the test.[85]

Part II: Contrast sensitivity tests

Implications for practice and future research

In the second part of this thesis the relation of spatial or temporal CS with reading was analysed. For a normally sighted group of participants, a strong correlation between spatial CS and reading speed was found and the influence on reading speed seems to be mainly optical in origin and, therefore, related to age effects. For subjects with macular pathologies the strong association between temporal CS and reading speed is suggested to reflect a high sensitivity for neural integrity of the temporal CS test. Results of the correlation between spatial CS and reading speed are in accordance with earlier research.[70,75,86] The association between temporal CS and reading speed has not earlier been described for participants with macular pathologies.

The temporal CS test can be considered a more basic test, as it addresses neural function, whereas VA and spatial CS reflect both neural and optical properties of the eye. It was found that the reliability and specificity of the short-duration temporal CS test used in this thesis for assessing neural integrity was good. A downside of the temporal CS test is that not all participants with severe maculopathies could manage to perform the test due to profound vision loss. One could speculate about an explanation for the high sensitivity of the temporal CS test for neural integrity: It is possible that spatial and temporal CS target different retinal mechanisms. Flicker stimulation (in comparison to a static stimulation) is reported to lead to a higher metabolic demand of retinal tissue.[87-89] In eyes with ARM, the retinal tissue is affected allowing flicker stimuli to detect (early) functional retinal deficiencies even before VA decreases.[90-92] It was therefore suggested in the literature that functional assessment of retinal integrity and retinal changes with a temporal CS test, in combination with imaging techniques and clinical grading, could have a high potential for diagnosing, monitoring and predicting disease progression in ARM.[73,74,92,93] It would be interesting to study whether the short-duration temporal CS test used in the present study could serve this purpose. In addition it would be interesting to study temporal CS and the relation with reading speed for patients with early maculopathies to provide more insight into patients with early ARM with adequate high contrast distance VA, but who experience reading problems.

Reading speed has also been proposed as a potential vision test to predict potential vision after cataract surgery.[72,94] However, reading is a complex task in which various patient characteristics (e.g. age, education, cognition) play a role, making it more difficult to differentiate between neural factors or patient characteristics in the case of low reading speeds. On the other hand, temporal CS tests might incidentally give unreliable results, as temporal CS seems to be insensitive to the presence of amblyopia.[71,95]

Both the Mars letter CS test and the temporal CS with the C-Quant device were shown to be a reliable test for measuring spatial CS and temporal CS, respectively.

The coefficient of reproducibility for spatial CS found in this thesis was in agreement with other studies.[96,97] Both studies used the Mars letter Contrast Sensitivity Test to measure spatial CS. In this thesis it was the first time that repeatability and reproducibility of temporal CS measured with the adaptation to the C-Quant device for a group of participants with maculopathies was described. The reliability of the temporal CS needs further research in visually impaired participants with other ocular pathologies of which early ARM is of particular interest. Considering the agreement between the spatial CS and temporal CS methods, the values of temporal CS are 0.1 log units higher than the spatial CS values. This is a small difference, and one can only speculate as to its significance and origin. It can potentially be attributed to optical mechanisms leading to a worsened image on the retina for the spatial (specifically, letter) CS. Because the retinal processing involved is different between the two tests, potentially a difference between the neurons in the parvo-cellular and magno-cellular pathways mediating primarily the spatial CS and temporal CS, respectively, can explain the difference.

Strengths and limitations

In this thesis, some new knowledge has been generated. The psychophysical properties of several reading tests were evaluated and some even for the first time (i.e. LEO and NED). Furthermore it is the first study to present repeatability and/or reproducibility data on the IReST and Colenbrander Reading Card in a group of visually impaired patients with maculopathies. In addition, the reliability of two different CS tests was evaluated. This thesis describes the repeatability and reproducibility of temporal CS measured with the adaptation to the C-Quant device for a group of participants with maculopathies, which is new as well.

There are some limitations to be discussed. All results are based on the Dutch version (a Germanic language) of the reading tests. It is possible that other results may emerge when using other languages. However for testing existing reading tests and the development of new tests in other languages the several recommendations made in this thesis can still be applied. A limitation of the systematic review is that there is no consensus on how to assess the quality of measurement properties of reading charts. Using the COSMIN checklist may be considered a limitation, because it was developed to assess patient reported outcome measures, such as self-report questionnaires. Therefore, we had to slightly adapt the criteria. It would be interesting to establish consensus among experts in the field of reading tests on the definitions of measurement properties and the criteria that should at least be assessed.

The reading tests studied in this thesis were all designed for reading out loud and are convenient for objective reading performance measurements. The rater can score the accuracy of words read and time taken to read words by listening. A disadvantage often mentioned of reading out loud is the possible ceiling effects that may occur when the participants speaking rate is slower than the visual decoding of text.[22]

However, for most visually impaired participants who have reading speeds below 200 WPM this is unlikely to be a limiting factor. Practically it is much more difficult to assess reading speed in case of silent reading. Comprehension tests are needed to ensure silently read text is accurately read, and not just skimmed. In addition, comprehension is less closely coupled to visual processing, it is difficult to measure accurately, and it is time consuming.[98] Furthermore, there is a gradual transition between good and poor comprehension and substantial inter-individual variability in comprehension performance due to cognitive factors. It was also shown that oral and silent reading are approximately the same for procedures where oral reading is used to check for accuracy of reading of short text.[99,100]

Another disadvantage often mentioned of continuous text reading tests is that due to contextual cues reading speed and acuity could be overestimated. Context plays an important role in experienced fluent readers, and contextual cues can encourage and assist guessing. Text which has a clear context increases reading speed in comparison with text containing unconnected random words.[57,101] The unpredictability of random word and letter sequences was suggested to make reading performance more dependent on eyesight and less dependent on reading skill and educational level. The type of reading errors can even indicate the presence and location of a scotoma.[102] On the other hand continuous text reading test are a more representative reading task for daily reading materials in comparison with unrelated word tests.[47]

Of the data presented in this thesis, the way scotomas were assessed might be considered a limitation. During the period the study was performed, no micro-perimeters, which includes a scanning laser ophthalmoscope, were available. Therefore, a clearly-defined size and location of the macular scotomata could not be determined. Further research using a micro-perimeter might provide more insight into the relation between CS and reading, and the influence of macular problems.[55] However, to address this problem, spatial CS, temporal CS, reading speed and the Amsler chart were assessed monocularly. To assess all tests monocularly, it was assumed that the preferred retinal locus of the participants would be the same for the different tests, as no shift could be made to the fellow eye. However, due to a possible unstable fixation, this cannot be guaranteed and could have lead to deficits in reading speed.[103]

CONCLUSION

The first part of this thesis shows that reading performance results obtained with function-based reading tests are not always reliable, and reading parameters cannot always be properly assessed. It would be beneficial to reconsider the use of certain reading tests in the ophthalmologic and optometric practice and research, because of unsatisfactory psychophysical properties. This was recently done by the Dutch association for occupational medicine regarding the directives of computer work, where recommendations were made to use the Radner Reading Charts for assessing near vision.[104] The choices regarding which reading test to use for clinical practice, and especially for research purposes, should be based on the reliability and feasibility of the reading test. In conclusion, the use of the Colenbrander Reading Card is debatable as the sentence equality is low and ceiling effects were found for a young and normally sighted population. The NED proved to be inappropriate for normally sighted subjects and subjects with maculopathies. It is recommended that the NED is no longer used in clinical practice and scientific research. The LEO had a low percentage of sentence equality, however the inter-session reliability was good and the agreement with the Radner Reading Charts was high. The LEO can be considered an option for the Dutch clinical practice, however it is not favourable for research. The Radner Reading Charts is the most favourable reading chart for clinical practice and research, as it has the highest percentage of sentence equality, good inter-session reliability and is found to be most reliable in determining the CPS for a group of patients with maculopathies. In addition, the Radner Reading Charts are available in several European languages, making it the test which is most suitable for research across different languages in order to allow standardised and comparable analysis of reading performance. Longer text paragraphs can be a supplemental tool to provide information on reading performance for prolonged reading. The longer paragraphs of the IReST seem promising, but need further reliability and feasibility studies, especially among visually impaired participants.

The second part of the thesis shows that there is significant correlation between spatial CS or temporal CS and reading speed. For a normally sighted group of participants, the strongest correlation was found between spatial CS and reading speed and is suggested to reflect a deterioration of both optical and neural factors with increasing age. For a group of participants with maculopathies, the strongest association was found between temporal CS and reading speed, and is suggested to reflect a high sensitivity for neural integrity of the temporal CS test. The temporal CS test used in the present study might be promising as a diagnostic test to detect and monitor the progression of early ARM and to assess the efficacy of interventions. Both the Mars Letter CS Test and the temporal CS with the C-Quant device were shown to be reliable for measuring spatial CS and temporal CS, respectively.

REFERENCES

1. Hazel CA, Petre KL, Armstrong RA, Benson MT, Frost NA. Visual function and subjective quality of life compared in subjects with acquired macular disease. Invest Ophthalmol Vis Sci 2000;41:1309-1315.
2. Bansback N, Czoski-Murray C, Carlton J, et al. Determinants of health related quality of life and health state utility in patients with age related macular degeneration: the association of contrast sensitivity and visual acuity. Qual Life Res 2007;16:533-543.
3. Chachamovich E, Fleck M, Laidlaw K, Power M. Impact of major depression and subsyndromal symptoms on quality of life and attitudes toward aging in an international sample of older adults. Gerontologist 2008;48:593-602.
4. van der Aa HP, Comijs HC, Penninx BW, van Rens GH, van Nispen RM. Major depressive and anxiety disorders in visually impaired older adults. Invest Ophthalmol Vis Sci 2015;56:849-854.
5. Bourne RRA, Flaxman SR, Braithwaite T, et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. Lancet Glob Health 2017;5:e888-e897.
6. Limburg H, Keunen JE. Blindness and low vision in The Netherlands from 2000 to 2020-modeling as a tool for focused intervention. Ophthalmic Epidemiol 2009;16:362-369.
7. Elliott DB, Trukolo-Ilic M, Strong JG, Pace R, Plotkin A, Bevers P. Demographic characteristics of the vision-disabled elderly. Invest Ophthalmol Vis Sci 1997;38:2566-2575.
8. Brown JC, Goldstein JE, Chan TL, Massof R, Ramulu P. Characterizing functional complaints in patients seeking outpatient low-vision services in the United States. Ophthalmology 2014;121:1655-1662.
9. Resnikoff S, Keys TU. Future trends in global blindness. Indian J Ophthalmol 2012;60:387-395.
10. Binns AM, Bunce C, Dickinson C, et al. How effective is low vision service provision? A systematic review. Surv Ophthalmol 2012;57:34-65.
11. Alio JL, Alio Del Barrio JL, Vega-Estrada A. Accommodative intraocular lenses: where are we and where we are going. Eye Vis (Lond) 2017;4:16.
12. Ahn SJ, Legge GE. Psychophysics of reading--XIII. Predictors of magnifier-aided reading speed in low vision. Vision Res 1995;35:1931-1938.
13. Stifter E, Sacu S, Benesch T, Weghaupt H. Impairment of visual acuity and reading performance and the relationship with cataract type and density. Invest Ophthalmol Vis Sci 2005;46:2071-2075.
14. Legge GE, Ross JA, Isenberg LM, LaMay JM. Psychophysics of reading. Clinical predictors of low-vision reading speed. Invest Ophthalmol Vis Sci 1992;33:677-687.
15. Ahn SJ, Legge GE, Luebker A. Printed cards for measuring low-vision reading speed. Vision Res 1995;35:1939-1944.
16. Burggraaff MC, van Nispen RM, Melis-Dankers BJ, van Rens GH. Effects of standard training in the use of closed-circuit televisions in visually impaired adults: design of a training protocol and a randomized controlled trial. BMC Health Serv Res 2010;10:62.
17. International Council of Ophthalmology, Visual Function Committee [Internet] . Visual Acuity Measurement Standard. Italian Journal of Ophthalmology 1988:1-15. [cited at 2013 October 8]. Available from: http://www.icoph.org/dynamic/attachments/resources/icovisualacuity1984.pdf
18. Verezen CA, Hardus P, Treffers WF. [How valid is the disability percentage?] Dialoog 1997;10:13-16.
19. Vos JJ, Padmos P, Boogaard J. Occupational testing of near vision. Ophthalmic Physiol Opt 1994;14:413-418.
20. Radner W. Near vision examination in presbyopia patients: Do we need good homologated near vision charts? Eye Vis (Lond) 2016;3:29.
21. Rubin GS. Measuring reading performance. Vision Res 2013.

8

22. Pelli DG, Bex P. Measuring contrast sensitivity. Vision Res 2013;90:10-14.
23. West CG, Gildengorin G, Haegerstrom-Portnoy G, Schneck ME, Lott L, Brabyn JA. Is vision function related to physical functional ability in older adults? J Am Geriatr Soc 2002;50:136-145.
24. Mones J, Rubin GS. Contrast sensitivity as an outcome measure in patients with subfoveal choroidal neovascularisation due to age-related macular degeneration. Eye (Lond) 2005;19:1142-1150.
25. West SK, Rubin GS, Broman AT, Munoz B, Bandeen-Roche K, Turano K. How does visual impairment affect performance on tasks of everyday life? The SEE Project. Salisbury Eye Evaluation. Arch Ophthalmol 2002;120:774-780.
26. National Research Council. Visual Impairments: Determining Eligibility for Social Security Benefits Washington, DC: National Academy Press; 2002:1-354.
27. Rubin GS, Legge GE. Psychophysics of reading. VI--The role of contrast in low vision. Vision Res 1989;29:79-91.
28. Arditi A. Improving the design of the letter contrast sensitivity test. Invest Ophthalmol Vis Sci 2005;46:2225-2229.
29. Thayaparan K, Crossland MD, Rubin GS. Clinical assessment of two new contrast sensitivity charts. Br J Ophthalmol 2007;91:749-752.
30. Mayer MJ, Spiegler SJ, Ward B, Glucs A, Kim CB. Foveal flicker sensitivity discriminates ARM-risk from healthy eyes. Invest Ophthalmol Vis Sci 1992;33:3143-3149.
31. Breton ME, Wilson TW, Wilson R, Spaeth GL, Krupin T. Temporal contrast sensitivity loss in primary open-angle glaucoma and glaucoma suspects. Invest Ophthalmol Vis Sci 1991;32:2931-2941.
32. Nakamura M, Yamamoto M. Variable pattern of visual recovery of Leber's hereditary optic neuropathy. Br J Ophthalmol 2000;84:534-535.
33. Felius J, Swanson WH. Photopic temporal processing in retinitis pigmentosa. Invest Ophthalmol Vis Sci 1999;40:2932-2944.
34. van den Berg TJ, Franssen L, Kruijt B, Coppens JE. Psychophysics, reliability, and norm values for temporal contrast sensitivity implemented on the two alternative forced choice C-Quant device. J Biomed Opt 2011;16:085004.
35. van den Berg TJ, Franssen L, Kruijt B, Coppens JE. History of ocular straylight measurement: A review. Z Med Phys 2013;23:6-20.
36. Franssen L, Coppens JE, van den Berg TJ. Compensation comparison method for assessment of retinal straylight. Invest Ophthalmol Vis Sci 2006;47:768-776.
37. Legge GE, Rubin GS, Luebker A. Psychophysics of reading--V. The role of contrast in normal vision. Vision Res 1987;27:1165-1177.
38. Whittaker SG, Lovie-Kitchin J. Visual requirements for reading. Optom Vis Sci 1993;70:54-65.
39. Akutsu H, Legge GE, Ross JA, Schuebel KJ. Psychophysics of reading--X. Effects of age-related changes in vision. J Gerontol 1991;46:325-331.
40. Leat SJ, Woo GC. The validity of current clinical tests of contrast sensitivity and their ability to predict reading speed in low vision. Eye (Lond) 1997;11:893-899..
41. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539-549.
42. Patel PJ, Chen FK, Rubin GS, Tufail A. Intersession repeatability of visual acuity scores in age-related macular degeneration. Invest Ophthalmol Vis Sci 2008;49:4347-4352.
43. Altmand DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. Statistician 1983:307-317.
44. McAlinden C, Khadka J, Pesudovs K. Statistical methods for conducting agreement (comparison of clinical tests) and precision (repeatability or reproducibility) studies in optometry and ophthalmology. Ophthalmic Physiol Opt 2011;31:330-338.
45. Mansfield JS, Legge GE. To the editor: Reliability testing of the Dutch version of the Radner Reading Charts. Optom Vis Sci 2008;85:1201-1202.

46. Maaijwee K, Radner W, Van Meurs JC. Author's response. Optom Vis Sci 2008;85:1202-1204.
47. Bailey IL, Lovie JE. The design and use of a new near-vision chart. Am J Optom Physiol Opt 1980;57:378-387.
48. Radner W, Obermayer W, Richter-Mueksch S, Willinger U, Velikay-Parel M, Eisenwort B. The validity and reliability of short German sentences for measuring reading speed. Graefes Arch Clin Exp Ophthalmol 2002;240:461-467.
49. Trauzettel-Klosinski S, Dietz K. Standardized assessment of reading performance: the New International Reading Speed Texts IReST. Invest Ophthalmol Vis Sci 2012;53:5452-5461.
50. Hahn GA, Penka D, Gehrlich C, et al. New standardised texts for assessing reading performance in four European languages. Br J Ophthalmol 2006;90:480-484.
51. Kooijman AC, Beerthuizen JJG. [The new reading charts, M-unit.]. In: Vision '96: Proceedings of the International Low Vision Conference. Book I. ONCE: Madrid,1996:pp.39-44.
52. Radner W, Willinger U, Obermayer W, Mudrich C, Velikay-Parel M, Eisenwort B. [A new reading chart for simultaneous determination of reading vision and reading speed]. Klin Monbl Augenheilkd 1998;213:174-181.
53. Carver RP. Reading Rate: A Review of Research and Theory, Academic Press: San Diego, CA, 1990.
54. Rubin GS, West SK, Munoz B, et al. A comprehensive assessment of visual impairment in a population of older Americans. The SEE Study. Salisbury Eye Evaluation Project. Invest Ophthalmol Vis Sci 1997;38:557-568.
55. Calabrese A, Bernard JB, Hoffart L, et al. Wet versus dry age-related macular degeneration in patients with central field loss: different effects on maximum reading speed. Invest Ophthalmol Vis Sci 2011;52:2417-2424.
56. Calabrese A, Bernard JB, Faure G, Hoffart L, Castet E. Eye movements and reading speed in macular disease: the shrinking perceptual span hypothesis requires and is supported by a mediation analysis. Invest Ophthalmol Vis Sci 2014;55:3638-3645.
57. Lott LA, Schneck ME, Haegerstrom-Portnoy G, Brabyn JA, Gildengorin GL, West CG. Reading performance in older adults with good acuity. Optom Vis Sci 2001;78:316-324.
58. Andersen GJ. Aging and Vision: Changes in Function and Performance from Optics to Perception. Wiley Interdiscip Rev Cogn Sci 2012;3:403-410.
59. Sass SM, Legge GE, Lee HW. Low-vision reading speed: influences of linguistic inference and aging. Optom Vis Sci 2006;83:166-177.
60. Owsley C. Aging and vision. Vision Res 2011;51:1610-1622.
61. Ramulu PY, Swenor BK, Jefferys JL, Rubin GS. Description and validation of a test to evaluate sustained silent reading. Invest Ophthalmol Vis Sci 2013;54:673-680.
62. Ramulu PY, Swenor BK, Jefferys JL, Friedman DS, Rubin GS. Difficulty with out-loud and silent reading in glaucoma. Invest Ophthalmol Vis Sci 2013;54:666-672.
63. Stangler-Zuschrott E. [Decreased reading speed and rapid fatigue as signs of disordered visual function]. Klin Monbl Augenheilkd 1990;196:150-157.
64. Bellmann C, Unnebrink K, Rubin GS, Miller D, Holz FG. Visual acuity and contrast sensitivity in patients with neovascular age-related macular degeneration. Results from the Radiation Therapy for Age-Related Macular Degeneration (RAD-) Study. Graefes Arch Clin Exp Ophthalmol 2003;241:968-974.
65. Owsley C, Sekuler R, Boldt C. Aging and low-contrast vision: face perception. Invest Ophthalmol Vis Sci 1981;21:362-365.
66. Wood J, Chaparro A, Hickson L. Interaction between visual status, driver age and distracters on daytime driving performance. Vision Res 2009;49:2225-2231.
67. Owsley C, McGwin G, Jr. Vision and driving. Vision Res 2010;50:2348-2361.
68. Woods RL, Strang NC, Atchison DA. Measuring contrast sensitivity with inappropriate optical correction. Ophthalmic Physiol Opt 2000;20:442-451.
69. Vianya-Estopa M, Douthwaite WA, Noble BA, Elliott DB. Capabilities of potential vision test measurements: clinical

evaluation in the presence of cataract or macular disease. J Cataract Refract Surg 2006;32:1151-1160.

70. Giacomelli G, Virgili G, Giansanti F, et al. Clinical and microperimetric predictors of reading speed in low vision patients: a structural equation modeling approach. Invest Ophthalmol Vis Sci 2013;54:4403-4408.
71. del Romo GB, Douthwaite WA, Elliott DB. Critical flicker frequency as a potential vision technique in the presence of cataracts. Invest Ophthalmol Vis Sci 2005;46:1107-1112.
72. Pesudovs K, Patel B, Bradbury JA, Elliott DB. Reading speed test for potential central vision measurement. Clin Experiment Ophthalmol 2002;30:183-186.
73. Dimitrov PN, Robman LD, Varsamidis M, et al. Visual function tests as potential biomarkers in age-related macular degeneration. Invest Ophthalmol Vis Sci 2011;52:9457-9469.
74. Mayer MJ, Ward B, Klein R, Talcott JB, Dougherty RF, Glucs A. Flicker sensitivity and fundus appearance in pre-exudative age-related maculopathy. Invest Ophthalmol Vis Sci 1994;35:1138-1149.
75. McClure ME, Hart PM, Jackson AJ, Stevenson MR, Chakravarthy U. Macular degeneration: do conventional measurements of impaired visual function equate with visual disability? Br J Ophthalmol 2000;84:244-250.
76. Brown B, Lovie-Kitchin J. Repeated visual acuity measurement: establishing the patient's own criterion for change. Optom Vis Sci 1993;70:45-53.
77. Lovie-Kitchin JE, Brown B. Repeatability and intercorrelations of standard vision tests as a function of age. Optom Vis Sci 2000;77:412-420.
78. Calabrese A, Owsley C, McGwin G, Legge GE. Development of a Reading Accessibility Index Using the MNREAD Acuity Chart. JAMA Ophthalmol 2016;134:398-405.
79. Vos Hd. [New reading chart designed by TNO] Ned Tijdschr Geneeskd 1969;113:909.
80. Radner W, Radner S, Diendorfer G. A new principle for the standardization of long paragraphs for reading speed analysis. Graefes Arch Clin Exp Ophthalmol 2016;254:177-184.
81. Rhiu S, Kim M, Kim JH, Lee HJ, Lim TH. Korean Version Self-testing Application for Reading Speed. Korean J Ophthalmol 2017;31:202-208.
82. Kingsnorth A, Wolffsohn JS. Mobile app reading speed test. Br J Ophthalmol 2015;99:536-539.
83. Dexl AK, Schlogel H, Wolfbauer M, Grabner G. Device for improving quantification of reading acuity and reading speed. J Refract Surg 2010;26:682-688.
84. Leung TW, Flitcroft DI, Wallman J, et al. A novel instrument for logging nearwork distance. Ophthalmic Physiol Opt 2011;31:137-144.
85. Calabrese A, Gamam S, Mansfield JS, Legge GE. Implementing the MNREAD Reading Acuity Test on an iPad3. Invest Ophthalmol Vis Sci 2014;55: ARVO Annual Meeting Abstract.
86. Cacho I, Dickinson CM, Smith HJ, Harper RA. Clinical impairment measures and reading performance in a large age-related macular degeneration group. Optom Vis Sci 2010;87:344-349.
87. Formaz F, Riva CE, Geiser M. Diffuse luminance flicker increases retinal vessel diameter in humans. Curr Eye Res 1997;16:1252-1257.
88. Garhofer G, Zawinka C, Resch H, Huemer KH, Dorner GT, Schmetterer L. Diffuse luminance flicker increases blood flow in major retinal arteries and veins. Vision Res 2004;44:833-838.
89. Shakoor A, Blair NP, Mori M, Shahidi M. Chorioretinal vascular oxygen tension changes in response to light flicker. Invest Ophthalmol Vis Sci 2006;47:4962-4965.
90. Neelam K, Nolan J, Chakravarthy U, Beatty S. Psychophysical function in age-related maculopathy. Surv Ophthalmol 2009;54:167-210.
91. Feigl B, Brown B, Lovie-Kitchin J, Swann P. Functional loss in early age-related maculopathy: the ischaemia postreceptoral hypothesis. Eye (Lond) 2007;21:689-696.
92. Phipps JA, Dang TM, Vingrys AJ, Guymer RH. Flicker perimetry losses in age-related macular degeneration. Invest Ophthalmol Vis Sci 2004;45:3355-3360.

93. Luu CD, Dimitrov PN, Robman L, et al. Role of flicker perimetry in predicting onset of late-stage age-related macular degeneration. Arch Ophthalmol 2012;130:690-699.
94. Patel B, Elliott DB, Whitaker D. Optimal reading speed in simulated cataract: development of a potential vision test. Ophthalmic Physiol Opt 2001;21:272-276.
95. Wildberger H, Junghardt A, Neetens A, van den Ende P. [Critical (foveal) flicker fusion frequency (CFF) is helpful in differential diagnosis of organic lesions from orthoptic amblyopias]. Klin Monbl Augenheilkd 1998;212:311-313.
96. Haymes SA, Roberts KF, Cruess AF, et al. The letter contrast sensitivity test: clinical evaluation of a new design. Invest Ophthalmol Vis Sci 2006;47:2739-2745.
97. Dougherty BE, Flom RE, Bullimore MA. An evaluation of the Mars Letter Contrast Sensitivity Test. Optom Vis Sci 2005;82:970-975.
98. Legge GE. Psychophysics of reading in normal and low vision. Lawrence Erlbaum Associates Publishers: Mahwah, NJ, 2007.
99. Legge GE, Pelli DG, Rubin GS, Schleske MM. Psychophysics of reading--I. Normal vision. Vision Res 1985;25:239-252.
100. Legge GE, Ross JA, Luebker A, LaMay JM. Psychophysics of reading. VIII. The Minnesota Low-Vision Reading Test. Optom Vis Sci 1989;66:843-853.
101. Fine EM, Hazel CA, Petre KL, Rubin GS. Are the benefits of sentence context different in central and peripheral vision? Optom Vis Sci 1999;76:764-769.
102. Mackeben M, Nair UK, Walker LL, Fletcher DC. Random word recognition chart helps scotoma assessment in low vision. Optom Vis Sci 2015;92:421-428.
103. Crossland MD, Culham LE, Rubin GS. Fixation stability and reading speed in patients with newly developed macular disease. Ophthalmic Physiol Opt 2004;24:327-333.
104. Verbeek JHAM, van der Beek AJ, Bank HJ, et al [Internet]. Dutch Association for Occupational Medicine. [Multidisciplinary Guideline Computerwork]: Utrecht 2013. [cited at 2017 October 2] Available from: https://www.nvab-online.nl/sites/default/files/bestanden-webpaginas/27-02-2014_RL_Computerwerk.pdf